`NEW ENGLAND COMMON ASSESSMENT PROGRAM (NECAP) FREQUENTLY ASKED QUESTIONS IN SCIENCE

View responses to questions proposed by Vermont teachers and administrators regarding the NECAP Science Assessment and its implications.

TEST DEVELOPMENT AND FORMAT

1. Who decides what is on the test?

- Measured Progress, our NECAP Science test development company, employs developers who are skilled in the content domains and the developmental level of students at grades 4, 8 and 11. These developers produce items aligned with Assessment Targets (see #5 below) formulated by science specialists from the Departments of Education in Vermont, New Hampshire and Rhode Island and aligned to Grade Expectations.
- The DOE science specialists then review the items and after several revisions the items are further evaluated by an Item Review Committee (IRC) composed of teachers from the three states.
- After another DOE science specialists' review of the IRC recommendations, the items are included as pilot items.
- Each item is then evaluated once more, based on the performance during the pilot test.
- Approved items become part of an item pool from which future NECAP Science Assessments are developed.

2. How many questions contribute to a student's score on the NECAP Science Assessment?

The NECAP Science Assessment at each grade level (Grade 4, Grade 8 and Grade 11) consists of 33 multiple choice items, 3 constructed response items, and 7 items contained in an Inquiry Task. These items total the 63 points from which scale scores are calculated.

3. How are the number and quality of released items determined?

Released items for the Science NECAP represent approximately 25% of the items in the assessment document. These items are selected by the state science specialists, because they represent significant content or analysis processes that we would like to bring to the attention of teachers. These are NOT, in any way, rejected items. Each year the Inquiry Tasks will be released in their entirety as part of these items.

4. When is the next administration of the NECAP Science Assessment?

The spring 2009 window for the Science NECAP Assessment is May 11-28, 2009.

TEST CONTENT

5. What content does the NECAP Science Assessment address?

The NECAP Science Assessment addresses four science content domains: Life Science, Physical Science, Earth and Space Science and Inquiry. Specific content Assessment Targets (found on the DOE Web site at http://education.vermont.gov/new/html/pgm_curriculum/science/necap_science.html) have been determined by the state science specialists. These targets are based on the National Science Education Standards and the American Association for the Advancement of Science Standards, as well as the Vermont Grade Expectations. A chart summarizing the

alignment of these Targets with the Vermont Grade Expectations can also be found on the DOE Web site (http://education.vermont.gov/new/html/pgm_curriculum/science/necap_science.html).

6. **Why does some of the vocabulary used in the assessment seem to represent 'trivia' rather than the big concepts of scientific literacy?**

   We recognize a concern about the specificity of the language in questions. Developers are always striving to strike a balance between concepts/big ideas and content examples that adequately represent those concepts. We will continue to work with the tri-state content specialists and developers to incorporate examples that may be more universally recognized.

SCORING AND PROFICIENCY LEVELS

7. **How are the tests scored?**

   Scoring of the NECAP Science Assessment is conducted by Measured Progress and is a critical portion of the service they provide. Scorers are trained extensively and supervised by lead scorers on site at the Measured Progress complex. Scorers are calibrated on every constructed response or Inquiry Task item separately. The scoring process involves a read-behind feature, where an individual scorer's work is evaluated closely, and if their evaluations do not measure up to the calibration standard, they are removed from scoring that item and retrained in the scoring process. Scorers have learned to expect this scrutiny of their work.

   The standards and rubrics that are used for the scoring of each item have been developed by Measured Progress lead scorers in conjunction with the state science specialists. Student responses that include correctly labeled scientific diagrams with minimal explanation are acceptable. We want this assessment to be a tool that is attainable for students and useful to teachers for science programs throughout Vermont.

8. **What are the categories of proficiency on the NECAP Science Assessment?**

   The proficiency categories on the NECAP Science Assessment are . . .
   - Proficient with Distinction
   - Proficient
   - Partially Proficient
   - Substantially Below Proficient

9. **How are the proficiency levels determined?**

   The points for proficiency at each level were determined by committees of teachers from VT, NH and RI during the Standard Setting process last August. Data for all Vermont schools are available on the DOE Web site at http://education.vermont.gov/new/html/pgm_assessment/performance/necap_science_public_schools_A_D.html.

For answers to additional questions you may have regarding the NECAP Science Assessment, call Measured Progress at this toll free Number: 1-877-632-7774.

IMPROVING STUDENT PERFORMANCE

10. Why was the student performance on the constructed response (CR) items generally so poor? What is the best way to prepare my students for the Constructed Response items of the NECAP Science Assessment?

Overall the state-wide performance on the constructed response items is not surprising. We have been aware that many students have difficulty expressing what they know about science. Science educators agree that clear communication is a critical part of what a scientist must do—and we support efforts for an improvement in those skills. One way to focus on determining a student's science understanding is to allow clearly labeled diagrams to suffice for a long paragraph as part of the CR items. Materials and strategies are also provided through Vermont Professional Development Network meetings to support teachers in helping students develop their written communication skills.

Here are a few questions to consider as educators help students express their science understandings:

- In answering CR-types of questions—How do your students actually *explain* their answers? Is the explanation complete and directed at what the question is asking?
- Do students understand the distinction between *explain* and *describe*? To *describe* is to convey a clear description of observations. To *explain* is much more extensive, including the why or the how as asked by the item, usually including a *because* statement with supporting evidence.

Our assessment history indicates that in all content areas scores tend to be low the first year and then rise steadily as schools increase efforts to align curriculum and classroom instruction with the grade expectations. We were also pleased to note that schools and supervisory unions/districts that are known to be further along in the alignment process tended to have relatively higher scores.

Actually, we were pleased that some schools did so much better than the state average. School personnel will have to look at the data re: Domains of Science (Life Science, Physical Science, Earth/Space Science and Inquiry) for their students to determine whether scores are related to content or to DOK (Depth of Knowledge) concern. The range of scores extended from 0 to 100% for individual items…and that includes all the ELL students and responses that could not be scored.

Generally:

1. Many students are not developing a deep understanding of science content. (That's not true for all schools. There may be inconsistencies of student understanding even within a school with outstanding science instruction!) The Inquiry Task data bear this out. The lowest scores (0 and 1) are on the DOK 3 questions. Look at the Item Analysis sheet for students in your school, and then see how they scored on each of the Inquiry Task items.

2. Students, overall, definitely need practice in expressing what they know via the written word. They need practice, practice and more practice on answering explanation kinds of questions. Most of the Inquiry Task involves written communication.

3. Students, overall, need practice in designing their own investigations, organizing their thoughts and setting up a testable investigation by intentionally controlling variables and planning multiple trials, and then conveying that information clearly.

4. Students still need more opportunities for thoughtful inquiry investigation to help them become more familiar with Inquiry–thinking and scientific problem-solving. It seems many students around the state (despite curriculum reform) still do not have strong inquiry skills.

5. Here are some questions for educators to ask themselves:
    - Do teachers use classroom assessment on an ongoing basis to check student thinking during a unit of study and then modify instruction accordingly?
    - Are there grade level assessments in place at the end of a unit to determine if students achieved the level of conceptual understanding that was intended?
    - Do teachers have the opportunity to discuss student work generated by common assessments so they gain a better understanding of what it means to achieve a standard as well as discuss ways to improve instruction?

11. What are successful districts doing that might contribute to their strong scores on the Vermont NECAP Science Assessment?

Department of Education personnel interviewed superintendents, curriculum coordinators and principals in schools, districts and supervisory unions who had the highest NECAP Science scores. Several common themes present in these schools emerged from these interviews. Successful schools evidenced the following:

- District/Supervisory Union Curriculum aligned with Vermont Science Grade Expectations;
- Leadership that has emphasized importance of science and provided time and resources for teachers to improve instructional skills and develop science instructional materials and classroom assessments;
- Time and resources allotted to development and improvement of science instructional skills and materials;
- District/Supervisory Union Science Committees made up of teacher representatives from all grade levels (presumably leading to coordinated and coherent science instruction across grade levels);
- Instructional methods and materials that emphasize inquiry science;
- Development of classroom assessments that focus on inquiry; and
- Integration of reading and writing instruction using science (and math) content.

12. What is the best way to prepare my students for the Inquiry Task portion of the Science Assessment?

- The Inquiry Task portion of the NECAP Science Assessment is designed for students to demonstrate their understanding of inquiry and problem-solving skills in science. The Vermont Grade Expectations 1-8 for all grades cluster address these skills. (See the DOE Web site: http://education.vermont.gov/new/pdfdoc/pubs/grade_expectations/science.pdf.)
- Another document that can guide classroom instruction in inquiry is the Guidelines for the Development of Science Inquiry Tasks. (See the DOE Web site: http://education.vermont.gov/new/html/pgm_curriculum/science/necap_science.html.) Also, the River Deep Learning Village Web site (http://rlv.education.vermont.gov/lv/admin/login.jsp) contains a variety of Vermont teacher-developed assessment tasks that can be used as part of classroom instruction. Sample Inquiry Tasks are available for Grades 4, 8, and 11 from the NECAP Science Pilot Assessment (May 2007) and the 2008 NECAP Science Assessment (May 2008). (See the DOE Web site: http://education.vermont.gov/new/html/pgm_curriculum/science/necap_science.html.)

13. How can the DOE communicate to School Boards and communities the real importance of NECAP results to counter the misconceptions of the public?

Department of Education personnel are very careful to provide accurate information to the media regarding NECAP assessment results. The media receive background information on the purpose and meaning of the assessment data. Information that was provided to the media is available at this Web site.

- The purpose of the NECAP Science Assessment is to determine the 'science literacy' of students as defined by the National Science Education Standards--*"Scientific literacy is the knowledge and understanding of scientific concepts and processes required for personal decision making, participation in civic and cultural affairs, and economic productivity.***"**
- *'School and district results are used by school to help improve curriculum and instruction.'* (**NECAP School Report)**
- Students' performance on the NECAP Science assessment is most effectively used for program evaluation over a period of time. *'Individual student results are used to support information gathered through classroom instruction and assessments.'* **Cited from NECAP School Report**
- Because our Science Grade Expectations are arranged in grade clusters, an individual student's score on the NECAP science assessment is one indication of their **cumulative science learning** in grades K-4 or K-8 or K-12. No single teacher is responsible for the proficiency level of any student or any group of students.
- It is important to note that the results of NECAP assessments are only one measure of academic progress. Program and course content adjustments should be determined on the basis of at least three points of data. NECAP Science results constitute only one point. Other data points might include local common assessment data, classroom assessment data, end of course assessment data and Surveys of Enacted Curriculum analysis (SEC) (http://seconline.wceruw.org/resources.asp) or contact Pat Fitzsimmons at the Department of Education (pat.fitzsimmons@state.vt.us).